

On the condensed density of the generalized eigenvalues of pencils of Hankel Gaussian random matrices and applications

Piero Barone *

Abstract

Pencils of Hankel matrices whose elements have a joint Gaussian distribution with nonzero mean and not identical covariance are considered. An approximation to the distribution of the squared modulus of their determinant is computed which allows to get a closed form approximation of the condensed density of the generalized eigenvalues of the pencils. Implications of this result for solving several moments problems are discussed and some numerical examples are provided.

Key words: random determinants; complex exponentials; complex moments problem; logarithmic potentials; *AMS classification:*15A52, 44A60

*Istituto per le Applicazioni del Calcolo "M. Picone", C.N.R., Viale del Policlinico 137, 00161 Rome, Italy; e-mail: barone@iac.rm.cnr.it; fax: 39-6-4404306

Introduction

Let us define the random Hankel matrices

$$U_0 = \begin{bmatrix} a_0 & a_1 & \dots & a_{p-1} \\ a_1 & a_2 & \dots & a_p \\ \cdot & \cdot & \dots & \cdot \\ a_{p-1} & a_p & \dots & a_{n-2} \end{bmatrix}, \quad U_1 = \begin{bmatrix} a_1 & a_2 & \dots & a_p \\ a_2 & a_3 & \dots & a_{p+1} \\ \cdot & \cdot & \dots & \cdot \\ a_p & a_{p+1} & \dots & a_{n-1} \end{bmatrix} \quad (1)$$

where $n = 2p$, $a_k = s_k + \epsilon_k$, $k = 0, 1, 2, \dots, n-1$, ϵ_k is a complex Gaussian, zero mean, white noise, with variance σ^2 and $s_k \in \mathcal{C}$.

Let us consider the generalized eigenvalues $\{\xi_j, j = 1, \dots, p\}$ of (U_1, U_0) i.e. the roots of the polynomial $P(z) = \det[U_1 - zU_0]$ and the associated condensed density $h(z)$, introduced in [8], which is the expected value of the (random) normalized counting measure on the zeros of $P(z)$ i.e.

$$h(z) = \frac{1}{p} E \left[\sum_{j=1}^p \delta(z - \xi_j) \right]$$

or, equivalently, for all Borel sets $A \subset \mathcal{C}$

$$\int_A h(z) dz = \frac{1}{p} \sum_{j=1}^p \text{Prob}(\xi_j \in A).$$

It can be proved that (see e.g. [1]) $h(z) = \frac{1}{4\pi} \Delta u(z)$ where Δ denotes the Laplacian operator with respect to x, y if $z = x + iy$ and $u(z) = \frac{1}{p} E \{\log(|P(z)|^2)\}$ is the corresponding logarithmic potential.

The condensed density $h(z)$ plays an important role for solving moment problems such as the trigonometric, the complex, the Hausdorff ones. It was shown in [9, 10, 6, 5, 7, 4, 2, 3] that all these problems can be reduced to the

complex exponentials approximation problem (CEAP), which can be stated as follows. Let us consider a uniformly sampled signal made up of a linear combination of complex exponentials

$$s_k = \sum_{j=1}^p c_j \xi_j^k. \quad (2)$$

where $c_j, \xi_j \in \mathcal{C}$. Let us assume to know an even number $n \geq 2p$ of noisy samples

$$a_k = s_k + \epsilon_k, \quad k = 0, 1, 2, \dots, n-1$$

where ϵ_k is a complex Gaussian, zero mean, white noise, with finite known variance σ^2 . We want to estimate p, c_j, ξ_j , $j = 1, \dots, p$, which is a well known ill-posed inverse problem. We notice that, in the noiseless case and when $n = 2p$, the parameters ξ_j are the generalized eigenvalues of the pencil (U_1, U_0) where now U_0 and U_1 are built as in (1) but starting from $\{s_k\}$.

From its definition it is evident that the condensed density provides information about the location in the complex plane of the generalized eigenvalues of (U_1, U_0) whose estimation is the most difficult part of CEAP. Unfortunately its computation is very difficult in general. In [7] a method to solve CEAP was proposed based on an approximation of the condensed density. An explicit expression of $h(z)$ proposed by Hammersley [8] when the coefficients of $P(z)$ are jointly Gaussian distributed was used. The second order statistics of these coefficients in the CEAP case were estimated by computing many Pade' approximants of different orders to the Z -transform of the data $\{a_k\}$. This last step was essential to realize the averaging that appears in the definition of $h(z)$, which is the key feature to make the condensed density a useful

tool for applications. In fact in the noiseless case $h(z)$ is a sum of Dirac δ distributions centered on the generalized eigenvalues while, when the signal is absent ($s_k = 0 \quad \forall k$), it was proved in [1] that if $z = re^{i\theta}$, the marginal condensed density $h^{(r)}(r)$ w.r. to r of the generalized eigenvalues is asymptotically in n a Dirac δ supported on the unit circle $\forall \sigma^2$. Moreover for finite n the the marginal condensed density w.r. to θ is uniformly distributed on $[-\pi, \pi]$. Therefore if the signal-to-noise ratio (e.g. $SNR = \frac{1}{\sigma} \min_{h=1,p} |c_h|$) is large enough $h(z)$ has local maxima in a neighbor of each ξ_j , $j = 1, \dots, p$ and this fact can be exploited to get good estimates of ξ_j . However usually we have only one realization of the discrete process $\{a_k\}$, hence we cannot estimate $h(z)$ by averaging. We then look for an approximation of $h(z)$ which can be well estimated by a single realization of $\{a_k\}$. The specific algebraic structure of CEAP will be taken into account and it will be shown that the noise contribution to $h(z)$ can be smoothed out to some extent simply acting on a parameter of the approximant.

The paper is organized as follows. In Section 1 some algebraic preliminaries are developed. In Section 2 the closed form approximation of $h(z)$ is defined. In Section 3 we show how to get a smooth estimate of $h(z)$ from the data by exploiting its closed form approximation. Finally in Section 4 some numerical examples are provided.

1 Preliminaries

Let us consider the pencil $F = U_1 - zU_0$. The problem can be reduced to real (random) variables by using the following result ([8, Theorem 5.1]):

Proposition 1 *If $F = U_1 - zU_0$, $z \in \mathcal{C}$ and $F = V_R + iV_I$, $V_R, V_I \in \mathbb{R}^{p \times p}$, then*

$$|\det(F)|^2 = \det(G)$$

where G is the real isomorph of F , i.e.

$$G = \mathcal{R}(F) = \begin{bmatrix} V_R & -V_I \\ V_I & V_R \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Let us consider the QR decomposition of G where $Q^T Q = Q Q^T = I_n$ where T denotes transposition, R is an upper triangular matrix and I_n is the identity matrix of order n . We then have

$$|\det[U_1 - zU_0]|^2 = \det(G) = |\det(G)| = |\det(QR)| = |\det(R)| = \prod_{k=1,n} |R_{kk}|.$$

We are therefore interested on the distribution of $|R_{kk}|, k = 1, \dots, n$. In order to perform the QR decomposition of the random matrix $G = [\underline{g}_1, \dots, \underline{g}_n]$ we make use of the Gram-Schmidt algorithm because it produces a triangular matrix R with positive elements. It is given by where $Q = [\underline{q}_1, \dots, \underline{q}_n]$.

For $k = 1$ we get

$$R_{11} = \sqrt{\underline{g}_1^T \underline{g}_1} = \sqrt{\underline{\alpha}_1^T Z^T Z \underline{\alpha}_1}$$

For $k = 1, \dots, n$

$\underline{w}_k = \underline{g}_k$

if $k > 1$ then

$R_{ik} = \underline{q}_i^T \underline{g}_k, \ i = 1, \dots, k-1$

$\underline{w}_k = \underline{w}_k - \sum_{i=1}^{k-1} R_{ik} \underline{q}_i$

end

$R_{kk} = \sqrt{\underline{w}_k^T \underline{w}_k}$

$\underline{q}_k = \frac{\underline{w}_k}{R_{kk}}$

end

where, denoting by \Re, \Im the real and imaginary part of a complex number,

$$\underline{\alpha}_1 = [\Re(a_0), \dots, \Re(a_p), \Im(a_0), \dots, \Im(a_p)]^T$$

and

$$Z = \begin{bmatrix} Z_R & -Z_I \\ Z_I & Z_R \end{bmatrix} \in \mathbb{R}^{n \times (n+2)}, \quad Z_R + iZ_I = \begin{bmatrix} -z & 1 & 0 & \dots & 0 \\ 0 & -z & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & -z & 1 \end{bmatrix} \in \mathbb{R}^{p \times (p+1)}.$$

In fact by using the property of the real isomorph

$$\mathcal{R}(A)[\underline{a}, \underline{b}]^T = [\Re(A\underline{c}), \Im(A\underline{c})]^T, \quad \underline{c} = \underline{a} + i\underline{b}$$

if $\underline{\alpha}_1 = [\underline{\alpha}_{1R}, \underline{\alpha}_{1I}]^T$ and \underline{e}_k is the k -th column of I_n , we have

$$\begin{aligned} \underline{g}_1 &= G[\underline{e}_1, \underline{0}]^T = \mathcal{R}(F)[\underline{e}_1, \underline{0}]^T = [\Re(F\underline{e}_1), \Im(F\underline{e}_1)]^T \\ &= [\Re[(Z_R + iZ_I)(\underline{\alpha}_{1R} + i\underline{\alpha}_{1I})], \Im[(Z_R + iZ_I)(\underline{\alpha}_{1R} + i\underline{\alpha}_{1I})]]^T \\ &= \mathcal{R}[(Z_R + iZ_I)] [\underline{\alpha}_{1R}, \underline{\alpha}_{1I}]^T = Z\underline{\alpha}_1. \end{aligned}$$

We notice that $\underline{\alpha}_1$ is a random Gaussian $2(p+1)$ vector with mean

$$E[\underline{\alpha}_1] = [\Re(s_1), \dots, \Re(s_{p+1}), \Im(s_1), \dots, \Im(s_{p+1})]^T$$

and covariance matrix $\Sigma = \frac{\sigma^2}{2} I_{2(p+1)}$. Therefore R_{11} is distributed as the square root of a quadratic form in normal non-central variables. This is a well known distribution (see e.g. [11]).

For $k = 2$ we have

$$\begin{aligned} R_{12} &= \underline{q}_1^T \underline{g}_2 = \frac{\underline{g}_1^T \underline{g}_2}{\sqrt{\underline{g}_1^T \underline{g}_1}} \\ \underline{w}_k &= \underline{g}_2 - R_{12} \underline{q}_1 = \underline{g}_2 - \frac{\underline{g}_1^T \underline{g}_2}{\underline{g}_1^T \underline{g}_1} \underline{g}_1 \end{aligned}$$

and

$$R_{22} = \sqrt{\underline{w}_2^T \underline{w}_2}.$$

Therefore R_{22} is no longer distributed as a quadratic form. However we notice that the vector \underline{w}_2 is a linear combination of the Gaussian vectors \underline{g}_1 and \underline{g}_2 with weights 1 and $\frac{\underline{g}_1^T \underline{g}_2}{\underline{g}_1^T \underline{g}_1}$. If this random weight were deterministic, R_{22} would be distributed as the square root of a quadratic form in normal non-central variables. We then consider the following approximation. We first notice that in the Gram-Schmidt algorithm, \underline{w}_k , $k = 1, \dots, n$ can be

computed as

$$\begin{aligned}\underline{w}_1 &= \underline{g}_1 \\ \underline{w}_k &= \underline{g}_k - \sum_{i=1}^{k-1} R_{ik} \underline{q}_i = \underline{g}_k - \sum_{i=1}^{k-1} W_{ik} \underline{w}_i, \quad W_{ik} = R_{ik}/R_{ii}, \quad k = 2, \dots, n\end{aligned}$$

We then apply the Gram-Schmidt algorithm to the matrix

$$S = E[G] = \begin{bmatrix} E[V_R] & E[V_I] \\ -E[V_I] & E[V_R] \end{bmatrix}$$

where

$$E[V_R] + iE[V_I] = E[F] = S_1 - zS_0$$

and S_0, S_1 are the Hankel matrices built from $\{s_k\} = \{E[a_k]\}$ in the same way as U_0 and U_1 are built from $\{a_k\}$. Let $S = \tilde{Q}\tilde{R}$ the QR decomposition of the deterministic matrix S , and consider the approximation of the vectors \underline{w}_k given by:

$$\begin{aligned}\tilde{\underline{w}}_1 &= \underline{g}_1 \\ \tilde{\underline{w}}_k &= \underline{g}_k - \sum_{i=1}^{k-1} \tilde{W}_{ik} \tilde{\underline{w}}_i, \quad \tilde{W}_{ik} = \tilde{R}_{ik}/\tilde{R}_{ii}, \quad k = 2, \dots, n\end{aligned}$$

It turns out that $\tilde{\underline{w}}_k$ are linear combinations of the Gaussian vectors $\underline{g}_1, \dots, \underline{g}_k$ with deterministic weights 1 and \tilde{W}_{ik} and we will prove in the next section that

$$R_{kk} \approx \sqrt{\tilde{\underline{w}}_k^T \tilde{\underline{w}}_k}$$

which is distributed as the square root of a quadratic form in normal non-central variables. In fact we have

$$\underline{g} = \begin{bmatrix} \underline{g}_1 \\ \underline{g}_2 \\ \vdots \\ \underline{g}_n \end{bmatrix} = \begin{bmatrix} I_n & 0 & \cdots & 0 \\ \tilde{W}_{12} \cdot I_n & I_n & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{W}_{1,n} \cdot I_n & \tilde{W}_{2,n} \cdot I_n & \cdots & I_n \end{bmatrix} \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \\ \vdots \\ \tilde{w}_n \end{bmatrix} = (\tilde{W} \otimes I_n) \underline{\tilde{w}} \quad (3)$$

where

$$\tilde{W} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \tilde{W}_{12} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{W}_{1,n} & \tilde{W}_{2,n} & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad \underline{g} = (I_n \otimes Z) \underline{\alpha}, \quad \text{and} \quad \underline{\alpha} = \begin{bmatrix} \underline{\alpha}_1 \\ \underline{\alpha}_2 \\ \vdots \\ \underline{\alpha}_n \end{bmatrix} \in \mathbb{R}^{n(n+2)}$$

with $\underline{\alpha}_k \in \mathbb{R}^{n+2}$ and

$$\underline{\alpha}_k = \begin{cases} [\Re(a_{k-1}), \dots, \Re(a_{k+p-1}), \Im(a_{k-1}), \dots, \Im(a_{k+p-1})]^T & \text{if } k = 1, \dots, p \\ [-\Im(a_{k-1}), \dots, -\Im(a_{k+p-1}), \Re(a_{k-1}), \dots, \Re(a_{k+p-1})]^T & \text{if } k = p+1, \dots, n \end{cases}$$

hence, if $M_k = \underline{e}_k \otimes I_n$ and \underline{e}_k is the k -th column of the identity matrix I_n ,

we have

$$\begin{aligned} R_{kk} &\approx \sqrt{\tilde{w}_k^T \tilde{w}_k} = \sqrt{\tilde{w}^T M_k M_k^T \tilde{w}} = \sqrt{\underline{g}^T (\tilde{W}^{-T} \otimes I_n) M_k M_k^T (\tilde{W}^{-1} \otimes I_n) \underline{g}} = \\ &= \sqrt{\underline{g}^T (\tilde{W}^{-T} \otimes I_n) (\underline{e}_k \underline{e}_k^T \otimes I_n) (\tilde{W}^{-1} \otimes I_n) \underline{g}} = \sqrt{\underline{g}^T (\tilde{W}^{-T} \underline{e}_k \underline{e}_k^T \tilde{W}^{-1} \otimes I_n) \underline{g}} = \\ &= \sqrt{\underline{\alpha}^T (I_n \otimes Z^T) (\tilde{W}^{-T} \underline{e}_k \underline{e}_k^T \tilde{W}^{-1} \otimes I_n) (I_n \otimes Z) \underline{\alpha}} = \sqrt{\underline{\alpha}^T (\tilde{W}^{-T} \underline{e}_k \underline{e}_k^T \tilde{W}^{-1} \otimes Z^T Z) \underline{\alpha}}. \end{aligned}$$

But we notice that if $\underline{a} = [\Re(a_0), \dots, \Re(a_{n-1}), \Im(a_0), \dots, \Im(a_{n-1})]^T$ then $\underline{\alpha} =$

$B \underline{a}$ where

$$B_k = [0, \dots, 0, \overbrace{I_{p+1}}^{k-1}, 0, \dots, 0] \in \mathbb{R}^{(p+1) \times n}$$

and

$$B = \begin{bmatrix} B_1^T & 0 & \cdot & \cdot & \cdot & \cdot & B_p^T & 0 & 0 & -B_1^T & \cdot & \cdot & \cdot & \cdot & 0 & -B_p^T \\ 0 & B_1^T & \cdot & \cdot & \cdot & \cdot & 0 & B_p^T & B_1^T & 0 & \cdot & \cdot & \cdot & \cdot & B_p^T & 0 \end{bmatrix}^T \in \mathbb{R}^{n(n+2) \times 2n}.$$

Summing up, if $A_k = B^T(\tilde{W}^{-T}\underline{e}_k\underline{e}_k^T\tilde{W}^{-1} \otimes Z^T Z)B \in \mathbb{R}^{2n \times 2n}$ then it will be shown that

$$R_{kk} \approx \sqrt{\underline{\alpha}^T(\tilde{W}^{-T}\underline{e}_k\underline{e}_k^T\tilde{W}^{-1} \otimes Z^T Z)\underline{\alpha}} = \sqrt{\underline{a}^T A_k \underline{a}}$$

where $\underline{a} \sim N(\underline{s}, \frac{\sigma^2}{2}I_{2n})$ and $\underline{s} = [\Re(s_0), \dots, \Re(s_{n-1}), \Im(s_0), \dots, \Im(s_{n-1})]^T$.

Moreover the following Proposition holds true

Proposition 2 *If $A_k = P_k^T D_k P_k$ is the spectral decomposition of $A_k \in \mathbb{R}^{2n \times 2n}$, then*

$$\underline{a}^T A_k \underline{a} = \tilde{\underline{a}}_k^T \Lambda_k \tilde{\underline{a}}_k,$$

where $D_k = \text{diag}[\Lambda_k, 0]$, $\Lambda_k \in \mathbb{R}^{n \times n}$, $\tilde{\underline{a}}_k = [I_n, 0_{n \times n}]P_k \underline{a}$, and $\Lambda_k > 0$.

Moreover $\tilde{\underline{a}}_k \sim N(\underline{\mu}_k, \frac{\sigma^2}{2}I_n)$ where $\underline{\mu}_k = [I_n, 0_{n \times n}]P_k \underline{s}$.

Proof. By construction we know that

$$\underline{a}^T A_k \underline{a} = \tilde{\underline{w}}_k^T \tilde{\underline{w}}_k = \sum_{h=1}^n \tilde{\underline{w}}_{kh}^2$$

and $A_k \geq 0$. Hence by the Sylvester's law of inertia $n = \text{rank}(A_k)$. Therefore the diagonal matrix D_k of the eigenvalues of A_k can be partitioned as $D_k = \text{diag}[\Lambda_k, 0]$ where $\Lambda_k > 0$. Letting $\underline{x} = P_k \underline{a}$ and $\underline{x}^T = [\tilde{\underline{a}}_k^T, \underline{y}^T]$, we have

$$\underline{a}^T A_k \underline{a} = \underline{a}^T P_k^T D_k P_k \underline{a} = \underline{a}^T P_k^T \begin{bmatrix} \Lambda_k & 0 \\ 0 & 0 \end{bmatrix} D_k P_k \underline{a} = \underline{x}^T \begin{bmatrix} \Lambda_k & 0 \\ 0 & 0 \end{bmatrix} \underline{x} = \tilde{\underline{a}}_k^T \Lambda_k \tilde{\underline{a}}_k.$$

The last part of the thesis follows by noticing that $\underline{\mu}_k = E[\tilde{a}_k] = [I_n, 0_{n \times n}] P_k E[\underline{a}]$ and

$$E[(\tilde{a}_k - E[\tilde{a}_k])(\tilde{a}_k - E[\tilde{a}_k])^T] = [I_n, 0_{n \times n}] P_k E[(\underline{a} - E[\underline{a}])(\underline{a} - E[\underline{a}])^T] P_k^T [I_n, 0_{n \times n}]^T = \frac{\sigma^2}{2} [I_n, 0_{n \times n}] P_k P_k^T [I_n, 0_{n \times n}]^T = \frac{\sigma^2}{2} I_n.$$

□

2 Closed form approximation of $h(z)$

In order to build a closed form approximation of the distribution of $u(z)$ we notice that we can rewrite $u(z)$ as

$$u(z) = \frac{1}{p} E \{ \log(|P(z)|^2) \} = \frac{1}{2p} E \{ \log(|\det[F]|^4) \} = \frac{1}{n} E \{ \log(|\det[G]|^2) \} = \frac{1}{n} \sum_{k=1}^n E [\log(R_{kk}^2)]$$

Hence we will consider in the following the distribution of R_{kk}^2 and its approximation by the quadratic form in normal variables $\underline{a}^T A_k \underline{a}$.

Lemma 1 For $k = 2, \dots, n$; $i = 1, \dots, k-1$ let be $W_{ik} = R_{ik}/R_{ii}$ where $G = QR$ is the QR decomposition of the random matrix G and $\tilde{W}_{ik} = \tilde{R}_{ik}/\tilde{R}_{ii}$ where $S = \tilde{Q}\tilde{R}$ is the QR decomposition of the matrix $S = E[G]$. Then, for $\sigma \rightarrow 0$,

$$E[(W_{ik} - \tilde{W}_{ik})^2] = o(\sigma).$$

Proof. Let us consider the rational function $f_{ik}(\underline{a}) = W_{ik}$ and its Taylor expansion around $\underline{s} = E[\underline{a}]$:

$$f_{ik}(\underline{a}) = f_{ik}(\underline{s}) + \sum_h \frac{\partial f_{ik}}{\partial \underline{a}_h} \Big|_{\underline{s}} (\underline{a}_h - \underline{s}_h) + \dots$$

which can be rewritten as

$$W_{ik} - \tilde{W}_{ik} = \sum_h \beta_h (\underline{a}_h - \underline{s}_h) + \dots$$

and, squaring and taking expectations, remembering that $\underline{a} \sim N(\underline{s}, \frac{\sigma^2}{2} I_{2n})$ and applying Holder's inequality we get

$$\begin{aligned} E[(W_{ik} - \tilde{W}_{ik})^2] &= E \left[\left(\sum_h \beta_h (\underline{a}_h - \underline{s}_h) \right)^2 \right] + \dots \\ &\leq \sum_{hk} E [|\beta_h \beta_k| (\underline{a}_h - \underline{s}_h)(\underline{a}_k - \underline{s}_k)] + \dots \\ &\leq \sum_{hk} (E [|\beta_h \beta_k|^2])^{1/2} (E [(\underline{a}_h - \underline{s}_h)^2 (\underline{a}_k - \underline{s}_k)^2])^{1/2} \\ &= \sum_{h \neq k} (E [|\beta_h \beta_k|^2])^{1/2} (E [(\underline{a}_h - \underline{s}_h)^2] E [(\underline{a}_k - \underline{s}_k)^2])^{1/2} \\ &\quad + \sum_h (E [|\beta_h|^4])^{1/2} (E [(\underline{a}_h - \underline{s}_h)^4])^{1/2} + \dots \\ &= \frac{\sigma^2}{2} \sum_{h \neq k} (E [|\beta_h \beta_k|^2])^{1/2} + 3\sigma^4 \sum_h (E [|\beta_h|^4])^{1/2} + \dots \end{aligned}$$

The thesis follows because the dropped terms are functions of $\{\sigma^{2h}, h > 2\}$.

□

Theorem 1 Let \underline{x} and $\tilde{\underline{x}}$ be the vectors obtained by stacking respectively the elements W_{ik} , and \tilde{W}_{ik} , for $\{k = 2, \dots, n; i = 1, \dots, k-1\}$. Let be $q_k(\underline{x}) =$

R_{kk}^2 where $G = QR$ is the QR decomposition of G , and $\tilde{q}_k(\tilde{x}) = \underline{a}^T A_k \underline{a}$ where $A_k = B^T (\tilde{W}^{-T} \underline{e}_k \underline{e}_k^T \tilde{W}^{-1} \otimes Z^T Z) B$. Then, for $\sigma \rightarrow 0$,

$$E[|q_k(\underline{x}) - \tilde{q}_k(\tilde{x})|] = o(\sqrt{\sigma}).$$

Proof. Let us consider the rational function $q_k(\underline{x})$ and its Taylor expansion around \tilde{x} :

$$q_k(\underline{x}) = q_k(\tilde{x}) + \sum_h \frac{\partial q_k}{\partial x_h} \Big|_{\underline{x}=\tilde{x}} (x_h - \tilde{x}_h) + \dots$$

But, by definition, $q_k(\tilde{x}) = \tilde{q}_k(\tilde{x})$, hence, squaring and taking expectations, we have

$$\begin{aligned} E[|q_k(\underline{x}) - \tilde{q}_k(\tilde{x})|] &= E\left[\left|\sum_h \beta_h(x_h - \tilde{x}_h)\right|\right] + \dots \\ &\leq \sum_h E[|\beta_h(x_h - \tilde{x}_h)|] + \dots \text{ by Minkowsky's inequality} \\ &\leq \sum_h \left(E[|\beta_h|^2]\right)^{1/2} \left(E[(x_h - \tilde{x}_h)^2]\right)^{1/2} + \dots \end{aligned}$$

by Holder's inequality. The thesis follows from Lemma 1. \square

Theorem 1 allows us to assume that $\{R_{kk}^2, k = 1, \dots, n\}$ are approximately distributed as a quadratic form in normal variables provided that the noise is small enough. A large literature exists about distribution of quadratic forms in normal variables (see e.g. [11] and references therein). Several series expansions are considered for the density of the quadratic form. However we will make use of the approximation proposed in [13] which makes it possible to get a closed form expression of $E[\log(R_{kk}^2)]$.

Theorem 2 *Let be $\tilde{\underline{a}}_k \sim N(\underline{\mu}_k, \frac{\sigma^2}{2}I_n)$ and let $q_k = \tilde{\underline{a}}_k^T \Lambda_k \tilde{\underline{a}}_k$ be a quadratic form with $\text{rank}(\Lambda_k) = n$. Then there exist positive constants $\alpha_k, \beta_k, \gamma_k$ such that the first three moments of $\beta_k(\chi_{\alpha_k}^2)^{\gamma_k}$ match those of q_k .*

Proof. Let us denote by $\nu_{hk}, h = 1, 2, 3$ the first three moments w.r. to the origin of q_k , which exist and whose explicit expression can be found in [11, Theorem 3.2b.2]. The corresponding moments of $\beta_k(\chi_{\alpha_k}^2)^{\gamma_k}$ are

$$\begin{aligned}\tilde{\nu}_{1k} &= 2\beta_k \frac{\Gamma(\gamma_k + \alpha_k/2)}{\Gamma(\alpha_k/2)} \\ \tilde{\nu}_{2k} &= (2\beta_k)^2 \frac{\Gamma(2\gamma_k + \alpha_k/2)}{\Gamma(\alpha_k/2)} \\ \tilde{\nu}_{3k} &= (2\beta_k)^3 \frac{\Gamma(3\gamma_k + \alpha_k/2)}{\Gamma(\alpha_k/2)}\end{aligned}$$

Then β_k is obtained by solving

$$2\beta_k \frac{\Gamma(\hat{\gamma}_k + \hat{\alpha}_k/2)}{\Gamma(\hat{\alpha}_k/2)} = \nu_{1k}, \text{ i.e. } \hat{\beta}_k = \frac{\nu_{1k} \Gamma(\hat{\alpha}_k/2)}{2\Gamma(\hat{\gamma}_k + \hat{\alpha}_k/2)}$$

where $\hat{\alpha}_k$ and $\hat{\gamma}_k$ are obtained by solving the system of equations

$$\begin{cases} \frac{\Gamma(\alpha_k/2)\Gamma(2\gamma_k + \alpha_k/2)}{[\Gamma(\gamma_k + \alpha_k/2)]^2} = \frac{\nu_{2k}}{\nu_{1k}^2} \\ \frac{[\Gamma(\alpha_k/2)]^2\Gamma(3\gamma_k + \alpha_k/2)}{[\Gamma(\gamma_k + \alpha_k/2)]^3} = \frac{\nu_{3k}}{\nu_{1k}^3} \end{cases}$$

which admits an unique positive solution. In fact the mapping

$$f(\alpha, \gamma) = \begin{bmatrix} \frac{\Gamma(\alpha/2)\Gamma(2\gamma + \alpha/2)}{[\Gamma(\gamma + \alpha/2)]^2} \\ \frac{[\Gamma(\alpha/2)]^2\Gamma(3\gamma + \alpha/2)}{[\Gamma(\gamma + \alpha/2)]^3} \end{bmatrix}$$

is continuous in $\mathbb{R}^+ \times \mathbb{R}^+$ and

$$\frac{\partial f}{\partial \gamma} = \begin{bmatrix} 2f_1(\alpha, \gamma) [\Psi(\alpha/2 + 2\gamma) - \Psi(\alpha/2 + \gamma)] \\ 3f_2(\alpha, \gamma) [\Psi(\alpha/2 + 3\gamma) - \Psi(\alpha/2 + \gamma)] \end{bmatrix} > 0$$

because $f(\alpha, \gamma) > 0$ and $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is monotonic increasing for $x \in \mathbb{R}^+$. Therefore $f_1(\alpha, \gamma)$ and $f_2(\alpha, \gamma)$ are monotonic increasing for each fixed $\alpha > 0$ as functions of γ . Moreover we notice that $f(\alpha, 0) = 1, \forall \alpha$ and $d_1 = \frac{\nu_{2k}}{\nu_{1k}^2} \geq 1, d_2 = \frac{\nu_{3k}}{\nu_{1k}^3} \geq 1$ by Holder inequality, taking into account the positivity of the r.v. q_k . Therefore $f(\alpha, \gamma)$ can assume every value in $[1, \infty)$. \square

We can now prove the main theorem

Theorem 3 *For $\sigma \rightarrow 0$, and approximating the distribution of $R_{kk}^2(z), k = 1, \dots, n$ by the density $\beta_k(z)(\chi_{\alpha_k(z)}^2)^{\gamma_k(z)}$, where $Q(z)R(z)$ is the QR decomposition of $G(z)$, which is the real isomorph of $U_1 - zU_0$, we have*

$$u(z) = \frac{2}{n} E\{\log(|\det[U_1 - zU_0]|^2)\} \approx \tilde{u}(z)$$

where

$$\tilde{u}(z) = \frac{1}{n} \sum_{k=1}^n [\log(2)\gamma_k(z) + \log[\beta_k(z)] + \gamma_k(z)\Psi[\alpha_k(z)/2]$$

and

$$h(z) \approx \frac{1}{4\pi} \Delta \tilde{u}(z). \quad (4)$$

Proof. By Proposition 1 the computation of $|\det[U_1 - zU_0]|^4$ is reduced to that of $R_{kk}^2, k = 1, \dots, n$, whose density, for each k , can be approximated for $\sigma \rightarrow 0$ by that of a quadratic form in normal, noncentral, variables by Theorem 1. By Proposition 2 this quadratic form is equivalent to a quadratic form with diagonal positive matrix whose density can be approximated, up to

the third moment, by the density $\beta_k(\chi_{\alpha_k}^2)^{\gamma_k}$ for suitable constants $\alpha_k, \beta_k, \gamma_k$ by Theorem 2. But, if q is a r.v. with density $\beta(\chi_\alpha^2)^\gamma$, we have

$$E[\log(q)] = \frac{1}{2^{\alpha/2}\Gamma(\alpha/2)} \int_0^\infty \log(\beta q^\gamma) q^{\alpha/2-1} e^{-q/2} dq = \log(2)\gamma + \log(\beta) + \gamma\Psi(\alpha/2).$$

□

3 Smooth estimate of the condensed density

We want to show now that we can exploit the closed form expression of the condensed density to smooth out the noise contribution to $h(z)$. This allows us to get a good estimate of p and $\xi_j, j = 1, \dots, p$ which is the nonlinear most difficult part of CEAP, from a single realization of the measured process $\{a_k\}$.

We first notice that if in Theorem 2 we look for an approximant which fits the first two moments of q_k then the solution is

$$\hat{\gamma}_k = 1, \quad \hat{\beta}_k = \frac{\nu_{2k} - \nu_{1k}^2}{2\nu_{1k}}, \quad \hat{\alpha}_k = \frac{2\nu_{1k}^2}{\nu_{2k} - \nu_{1k}^2}$$

and $E[R_{kk}^2] = \nu_{1k} = \hat{\alpha}_k \hat{\beta}_k$. Hence, dropping the terms independent of z and using the approximation $\Psi(x) \approx \log(x) - \frac{1}{x}$ (see[12]) we get

$$\begin{aligned} h(z) &\approx \frac{1}{4\pi n} \sum_{k=1}^n \Delta (\log[\beta_k(z)] + \Psi[\alpha_k(z)/2]) \\ &\approx \frac{1}{4\pi n} \sum_{k=1}^n \Delta \left(\log[\beta_k(z)] + \log[\alpha_k(z)/2] - \frac{2}{\alpha_k(z)} \right) \\ &= \frac{1}{4\pi n} \sum_{k=1}^n \Delta \left(\log[\nu_{1k}(z)/2] - \frac{2}{\alpha_k(z)} \right). \end{aligned} \tag{5}$$

Now we notice that

$$\text{var}_{R_{kk}^2} = \nu_{2k} - \nu_{1k}^2 = \frac{\hat{\alpha}_k \hat{\beta}_k^2}{2} \text{ and } \frac{E[R_{kk}^2]}{\sqrt{\text{var}_{R_{kk}^2}}} = \sqrt{\frac{\hat{\alpha}_k}{2}} \quad (6)$$

Moreover from [11, Theorem 3.2b.2] the first two moments w.r. to the origin of R_{kk}^2 are

$$\begin{aligned} \nu_{1k} &= \frac{\sigma^2}{2} \sum_{h=1}^n \lambda_k(h) + \underline{\mu}_k^T \Lambda_k \underline{\mu}_k \\ \nu_{2k} &= \frac{\sigma^4}{2} \sum_{h=1}^n \lambda_k(h)^2 + 2\sigma^2 \underline{\mu}_k^T \Lambda_k^2 \underline{\mu}_k + \nu_{1k}^2 \end{aligned}$$

where $\lambda_k(h)$ are independent of σ^2 because they are the (positive) eigenvalues of A_k which is independent of σ^2 . Therefore when $\underline{s} \neq \underline{0}$ if

$$a = \frac{1}{2} \sum_{h=1}^n \lambda_k(h)^2 > 0, \quad b = \underline{\mu}_k^T \Lambda_k^2 \underline{\mu}_k > 0, \quad c = \sum_{h=1}^n \lambda_k(h) > 0, \quad d = 2\underline{\mu}_k^T \Lambda_k \underline{\mu}_k > 0$$

$$\hat{\beta}_k = \frac{\nu_{2k} - \nu_{1k}^2}{2\nu_{1k}} = \sigma^2 \left[\frac{\sigma^2 \frac{a}{2} + 2b}{\sigma^2 c + 2d} \right], \quad \hat{\alpha}_k = \frac{2\nu_{1k}^2}{\nu_{2k} - \nu_{1k}^2} = \frac{2 \left(\frac{\sigma^2}{2} c + d \right)^2}{\frac{\sigma^4 a}{2} + 2\sigma^2 b}$$

Deriving these expressions respectively with respect to σ^2 and to the SNR measured e.g. by $\rho = \frac{d}{\sigma^2}$ we get

$$\frac{\partial \hat{\beta}_k}{\partial \sigma^2} = \frac{bd + a\sigma^2(2d + c\sigma^2)}{(d + c\sigma^2)^2} > 0, \quad \frac{\partial \hat{\alpha}_k}{\partial \rho} = \frac{d(c + \rho)(2ad + b(\rho - c))}{(ad + b\rho)^2}.$$

But

$$\rho - c = \frac{d}{\sigma^2} - c = \frac{\underline{\mu}_k^T \Lambda_k \underline{\mu}_k}{\sigma^2/2} - \sum_{h=1}^n \lambda_k(h) = \sum_{h=1}^n \left(\frac{\mu_h^2}{\sigma^2/2} - 1 \right) \lambda_k(h) > 0$$

if $\mu_h^2 > \frac{\sigma^2}{2} \forall h$. Hence $\hat{\beta}_k$ is an increasing function of σ^2 and $\hat{\alpha}_k$ is an increasing function of ρ if the SNR is not too small consistently with the expression (6).

When $\sigma^2 = 0$

$$\hat{\beta}_k = 0, \quad \hat{\alpha}_k \hat{\beta}_k = \nu_{1k} = \underline{\mu}_k^T \Lambda_k \underline{\mu}_k, \quad \lim_{\sigma^2 \rightarrow 0} \hat{\alpha}_k = \infty$$

and, by (5)

$$\lim_{\sigma^2 \rightarrow 0} h(z) = \frac{1}{4\pi n} \sum_{k=1}^n \Delta(\log[\nu_{1k}(z)/2]).$$

The last equality can be checked also by noticing that

$$\nu_{1k}(z) = \underline{\mu}_k^T \Lambda_k \underline{\mu}_k = \underline{s}^H A_k \underline{s} = \tilde{R}_{kk}^2.$$

The idea is then to use the parameters $\hat{\beta}_k$ as smoothing parameters and $\hat{\alpha}_k$ as signal-related parameters. By fixing $\hat{\beta}_k = \sigma^2 \beta$, $\forall k$ and taking $\hat{\alpha}_k = \frac{\nu_{1k}}{\sigma^2 \beta}$ the variance of $R_{kk}^2(z)$ is controlled by β and $h(z)$ can be estimated by

$$\hat{h}(z) \propto \sum_{k=1}^n \hat{\Delta} \left(\Psi \left[\frac{\hat{\nu}_{1k}(z)}{2\sigma^2 \beta} \right] \right) \quad (7)$$

where $\hat{\Delta}$ is the discrete Laplacian and $\hat{\nu}_{1k}(z)$ is an estimate of $\nu_{1k}(z)$. In the following we use, as an estimate of $\nu_{1k}(z)$, the value $\hat{R}_{kk}^2(z)$ obtained by the QR decomposition of the data matrix $G(z)$ defined in Proposition 1.

From a qualitative point of view increasing β has the effect to make larger the support of all modes of $h(z)$ and to lower their value because $h(z)$ is a probability density. Hence the noise-related modes are likely to be smoothed out by a sufficiently large β . However a value of β too large can result in a low resolution spectral estimate.

4 Numerical results

In this section some experimental evidence of the claims made in the previous sections is given.

To appreciate the advantage of the closed form estimate $\hat{h}(z)$ with respect to an estimate of the condensed density obtained by MonteCarlo simulation an experiment was performed. $N = 100$ independent realizations of the r.v. $a_k^{(r)}, k = 1, \dots, n, \quad r = 1, \dots, N$ were generated from the complex exponentials model with $p = 5$ components given by

$$\underline{\xi} = [e^{-0.1-i2\pi 0.3}, e^{-0.05-i2\pi 0.28}, e^{-0.0001+i2\pi 0.2}, e^{-0.0001+i2\pi 0.21}, e^{-0.3-i2\pi 0.35}]$$

$$\underline{c} = [6, 3, 1, 1, 20], \quad n = 74.$$

We notice that the frequencies of the 3rd and 4th components are closer than the Nyquist frequency ($0.21 - 0.20 = 0.01 < 1/n = 0.0135$). Hence a super-resolution problem is involved in this case. Two values of the noise s.d. σ were used

$$\sigma = 0.2, \quad 0.8.$$

An estimate of $h(z)$ was computed on a square lattice of dimension $m = 100$ by

$$\hat{h}(z) \propto \sum_{r=1}^N \sum_{k=1}^n \hat{\Delta} \left(\Psi \left[\frac{R_{kk}^{(r)}(z)^2}{2\sigma^2\beta} \right] \right)$$

where $R^{(r)}(z)$ is obtained by the QR factorization of the matrix $G(z)$ built from sample r -th. In the top part of fig.1 the estimate of $h(z)$ obtained by Monte Carlo simulation is plotted. In the bottom part the smoothed estimates $\hat{h}(z)$ for $\sigma = 0.2$ and $\beta = 5n$ based on a single realization was plotted. In fig.2 the results obtained with $\sigma = 0.8$ and $\beta = 5n$ are shown. We notice that by the proposed method we get an improved qualitative information with respect to that obtained by replicated measures. This is an important

feature for applications where usually only one data set is measured. We also notice that when $\sigma = 0.2$ the probability to find a root of $P(z)$ in a neighbor of ξ_j is larger than the probability to find it elsewhere. This is no longer true when $\sigma = 0.8$ even if the signal-related complex exponentials are well separated. In the following we will say that the complex exponential model is identifiable if this last case occurs and it is strongly identifiable if the first case occurs. Therefore if the model is identifiable the signal-related complex exponentials are well separated but the relative importance of some of them - measured by the value of the local maxima of $h(z)$ - is not larger than the relative importance of some noise-related complex exponentials. Therefore in this case we need some a-priori information about the location of the ξ_j in order to separate signal-related components from the noise-related ones.

We want now to show by means of a small simulation study the quality of the estimates of the parameters $p, \underline{\xi}$ and \underline{c} which can be obtained from $\hat{h}(z)$. To this aim the following estimation procedure was used:

- the local maxima of $\hat{h}(z)$ are computed and sorted in decreasing magnitude
- a clustering method is used to group the local maxima into two groups. If the model is strongly identifiable the signal-related maxima are larger than the noise-related ones, therefore a simple thresholding is enough to separate the two groups. A good threshold is the one that produces an estimate of s_k which best fits the data a_k in L_2 norm as the noise is assumed to be Gaussian

- the cardinality \hat{p} of the class with largest average value is an estimate of p
- the local maxima $\hat{\xi}_j, j = 1, \dots, \hat{p}$ of the class with largest average value are estimates of $\xi_j, j = 1, \dots, p$. Of course if $\hat{p} \neq p$ some ξ_j are not estimated or viceversa some spurious complex exponentials are found
- \underline{c} is estimated by solving the linear least squares problem

$$\hat{\underline{c}} = \operatorname{argmin}_{\underline{x}} \|V\underline{x} - \underline{d}\|_2^2$$

where $V \in \mathbb{C}^{n \times \hat{p}}$ is the Vandermonde matrix based on $\hat{\xi}_j, j = 1, \dots, \hat{p}$

The bias, variance and mean squared error (MSE) of each parameter separately were estimated. $N = 500$ independent data sets $\underline{a}^{(r)}$ of length n were generated by using the model parameters given above and $\sigma = 0.2$. For $r = 1, \dots, N$ the condensed density estimate $\hat{h}^{(r)}(z)$ was computed on a square lattice of dimension $m = 100$. The estimation procedure is then applied to each of the $\hat{h}^{(r)}(z), r = 1, \dots, N$ and the corresponding estimates $\hat{\xi}_j^{(r)}, \hat{c}_j^{(r)}, j = 1, \dots, \hat{p}^{(r)}$ of the unknown parameters were obtained. If the estimate $\hat{p}^{(r)}$ was less than the true value p , the corresponding data set $\underline{a}^{(r)}$ was discarded.

In Table 1 the bias, variance and MSE of each parameter including p is reported. They were computed by choosing among the $\hat{\xi}_j^{(r)}, j = 1, \dots, \hat{p}^{(r)} \geq p$ the one closest to each $\xi_k, k = 1, \dots, p$ and the corresponding $\hat{c}_j^{(r)}$. If more than one ξ_k is estimated by the same $\hat{\xi}_j^{(r)}$ the r -th data set $\underline{a}^{(r)}$ was discarded. In the case considered all the data sets were accepted.

As a second example the reconstruction of a piecewise constant function from noisy Fourier coefficients is considered. The problem is stated as follows. Given a real interval $[-\pi, \pi]$ and $N+1$ numbers $-\pi \leq l_1 < l_2 < \dots < l_{N+1} \leq \pi$, let \mathcal{F} be the class of functions defined as

$$F(t) = \sum_{j=1}^N w_j \chi_j(t) ,$$

where

$$\chi_j(t) = \begin{cases} 1 & \text{if } t \in [l_j, l_{j+1}] \\ 0 & \text{otherwise} \end{cases} ,$$

and the w_j are real weights. The problem consists in reconstructing a function $F(t) \in \mathcal{F}$ from a finite number of its noisy Fourier coefficients

$$a_k = \frac{1}{2} \int_{-\pi}^{\pi} F(t) e^{itk} dt + \epsilon_k = s_k + \epsilon_k , \quad k = 0, \dots, n-1 ,$$

where ϵ_k is a complex Gaussian, zero mean, white noise, with variance σ^2 . We are looking for a solution which is not affected by Gibbs artifact and can cope, stably, with the noise. The basic observation is the following. The unperturbed moments s_k are given by

$$s_k = \frac{1}{2} \int_{-\pi}^{\pi} F(t) e^{itk} dt = \sum_{j=1}^N w_j \frac{\sin(\beta_j k)}{k} \exp(i\lambda_j k) ,$$

where

$$\beta_j = \frac{l_{j+1} - l_j}{2}, \quad \lambda_j = \frac{l_{j+1} + l_j}{2} .$$

Then consider the Z-transform of the sequence $\{s_k\}$

$$s(z) = \sum_{j=1}^N w_j \left(\beta_j + \frac{1}{2i} \ln \frac{z - e^{il_j}}{z - e^{il_{j+1}}} \right)$$

which converges if $|z| > 1$ and is defined by analytic continuation if $|z| \leq 1$. We notice that $s(z)$ has a branch point at $\xi_j = e^{il_j}$, $j = 1, \dots, N+1$ where l_j are the discontinuity points of $F(t)$. It was proved in [9, 10] that the c_j are strong attractors of the poles of the Pade' approximants $[m, n]_f(z)$ to the noisy Z -transform

$$f(z) = \sum_{k=0}^{\infty} a_k z^{-k}$$

when $m, n \rightarrow \infty$ and $m/n \rightarrow 1$. It is easy to show that the poles of $[m, n]_f(z)$ are the generalized eigenvalues of the pencil (U_1, U_0) built from the data a_k , $k = 0, \dots, n-1$ whose condensed density is $h(z)$. Therefore, as shown in [9, 10] the local maxima of $h(z)$ are concentrated along a set of arcs which ends in the branch points ξ_j and on a set of arcs close to the unit circle. As the branch points are strong attractors for the Pade' poles, the probability to find a pole in a neighbor of a branch point is larger than elsewhere, therefore it can be expected that the branch points correspond to the largest local maxima of $h(z)$, as far as the SNR is sufficiently large. In order to compute estimates \hat{l}_j of l_j , it is sufficient to compute the arguments of the main local maxima of $\hat{h}(z)$. The w_j are then estimated by taking the median of the data in each interval $[\hat{l}_j, \hat{l}_{j+1}]$. The median is in fact robust with respect to errors affecting \hat{l}_j .

The method was applied to an example considered in [10] where comparisons with other methods were also reported. In the top left part of fig.3 the original function $F(t)$ is plotted. In the top right the noisy data with $SNR = 7$ are reported where the SNR is measured as the ratio of the stan-

dard deviations of $\{s_k\}$ and $\{\epsilon_k\}$. In the bottom parts the condensed density and the reconstructed function $\hat{F}(t)$ are plotted. Looking at the condensed density we notice that the model is strongly identifiable, therefore the estimation procedure outlined above was applied. In fig.4 the same quantities as above but with $SNR = 1$ are plotted. In this case the model is identifiable but not strongly therefore the clustering step does not work. The number of complex exponentials used to get the reconstruction plotted in fig.4 is $\hat{p} = 20$ and was found by trial and errors.

We notice that when $SNR = 7$ we get an almost perfect reconstruction, better than that reported in [10]. When $SNR = 1$ the reconstruction quality is worse as expected but still comparable with the one reported in [10].

References

- [1] BARONE, P. (2005). On the distribution of poles of Pade' approximants to the Z-transform of complex Gaussian white noise, *J. Approx. Theory* **132** 224-240.
- [2] BARONE, P. (2003). Orthogonal polynomials, random matrices and the numerical inversion of Laplace transform of positive functions. *J.Comp. Applied Math.*, vol.155 **2** 307-330.
- [3] BARONE, P. (2003). Random matrices in Magnetic Resonance signal processing. The 8-th SIAM Conference on Applied Linear Algebra

- [4] BARONE, P., RAMPONI, A., SEBASTIANI, G.(2001). On the numerical inversion of the Laplace transform for Nuclear Magnetic Resonance relaxometry. *Inverse Problems* **17** 77-94.
- [5] BARONE, P., MARCH, R. (2001). A novel class of Padé based method in spectral analysis. *J. Comput. Methods Sci. Eng.* **1** 185-211.
- [6] BARONE, P., MARCH, R. (1998). Some properties of the asymptotic location of poles of Padé approximants to noisy rational functions, relevant for modal analysis. *IEEE Trans. Signal Process.* **46** 2448-2457.
- [7] BARONE, P., RAMPONI, A.(2000). A new estimation method in modal analysis. *IEEE Trans. Signal Process.* **48** 1002-1014.
- [8] J.M. HAMMERSLEY, *The zeros of a random polynomial*, Proc. Berkely Symp. Math. Stat. Probability, 3rd, 2 (1956), pp. 89–111.
- [9] MARCH, R., BARONE, P.(1998). Application of the Padé method to solve the noisy trigonometric moment problem: some initial results. *SIAM J. Appl. Math.* **58** 324-343.
- [10] MARCH, R., BARONE, P.(2000). Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method. *SIAM J. Appl. Math.* **60** 1137-1156.
- [11] MATHAI, A.M., PROVOST, B.(1977). *Quadratic forms in random variables*, Marcel Dekker, New York.

- [12] MUQATTASH, I., YADHI, M.(2006). Infinite family of approximations of the digamma function, *Mathematical and Computer Modelling***43** 1329-1336.
- [13] SOLOMON, H., STEPHENS, M.(1977). Distribution of a sum of weighted chi-square variables, *J. Am. Stat. Ass.***72** 881-885.

	p	$bias(\hat{p})$	$s.d.(\hat{p})$	$MSE(\hat{p})$
	5	0.0000	0.0000	0.0000
	ξ_j	$bias(\hat{\xi}_j)$	$s.d.\hat{\xi}_j$	$MSE(\hat{\xi}_j)$
$j = 1$	-0.2796 - 0.8606i	-0.0008 + 0.0001i	0.0000	0.0000
$j = 2$	-0.1782 - 0.9344i	0.0036 - 0.0010i	0.0000	0.0000
$j = 3$	0.3090 + 0.9510i	0.0057 - 0.0064i	0.0031	0.0001
$j = 4$	0.2487 + 0.9685i	-0.0058 + 0.0110	0.0019	0.0002
$j = 5$	-0.4354 + 0.5993i	-0.0047 + 0.0054i	0.0108	0.0002
	c_j	$bias(\hat{c}_j)$	$s.d.(\hat{c}_j)$	$MSE(\hat{c}_j)$
$j = 1$	6.0000	0.0440	0.1238	0.0173
$j = 2$	3.0000	-0.0407	0.0688	0.0064
$j = 3$	1.0000	0.0441	0.0736	0.0074
$j = 4$	1.0000	-0.6767	0.0808	0.4644
$j = 5$	20.0000	-0.1007	0.2574	0.0764

Table 1: Statistics of the parameters \hat{p} , $\hat{\xi}_j, j = 1, \dots, p$ and $\hat{c}_j, j = 1, \dots, p$

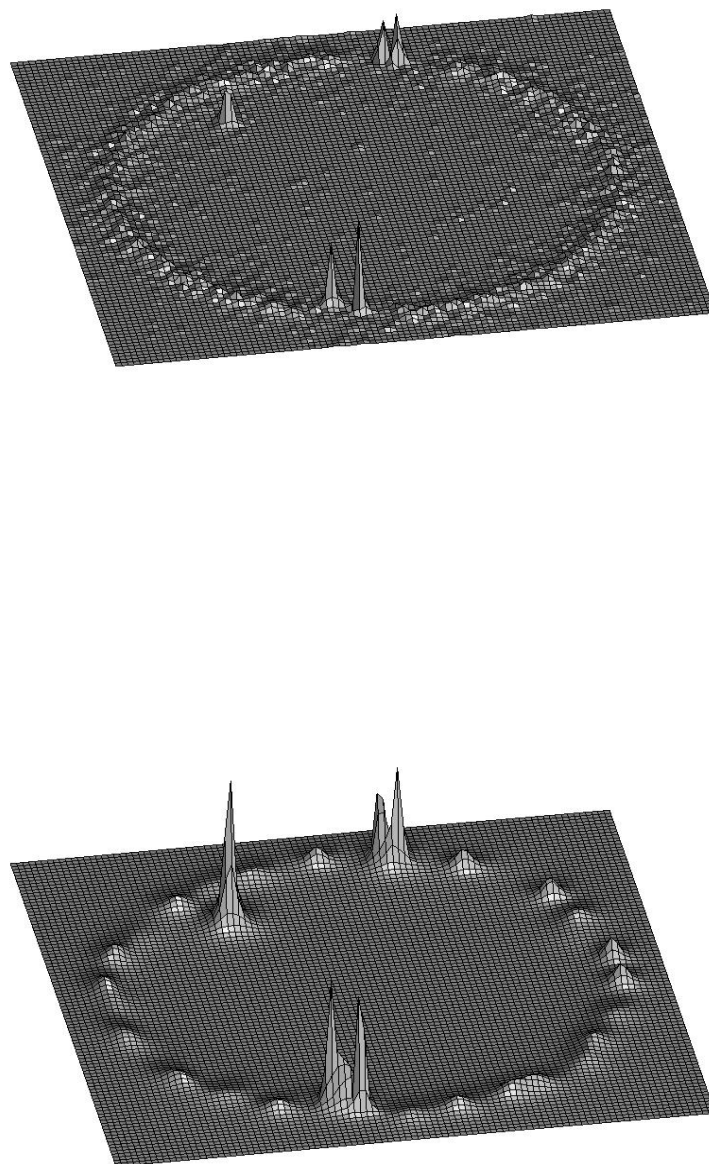


Figure 1: Top: Monte Carlo estimate of the condensed density when $\sigma = 0.2$; bottom: estimate of the condensed density by the closed form approximation with $\beta = 14.8$.

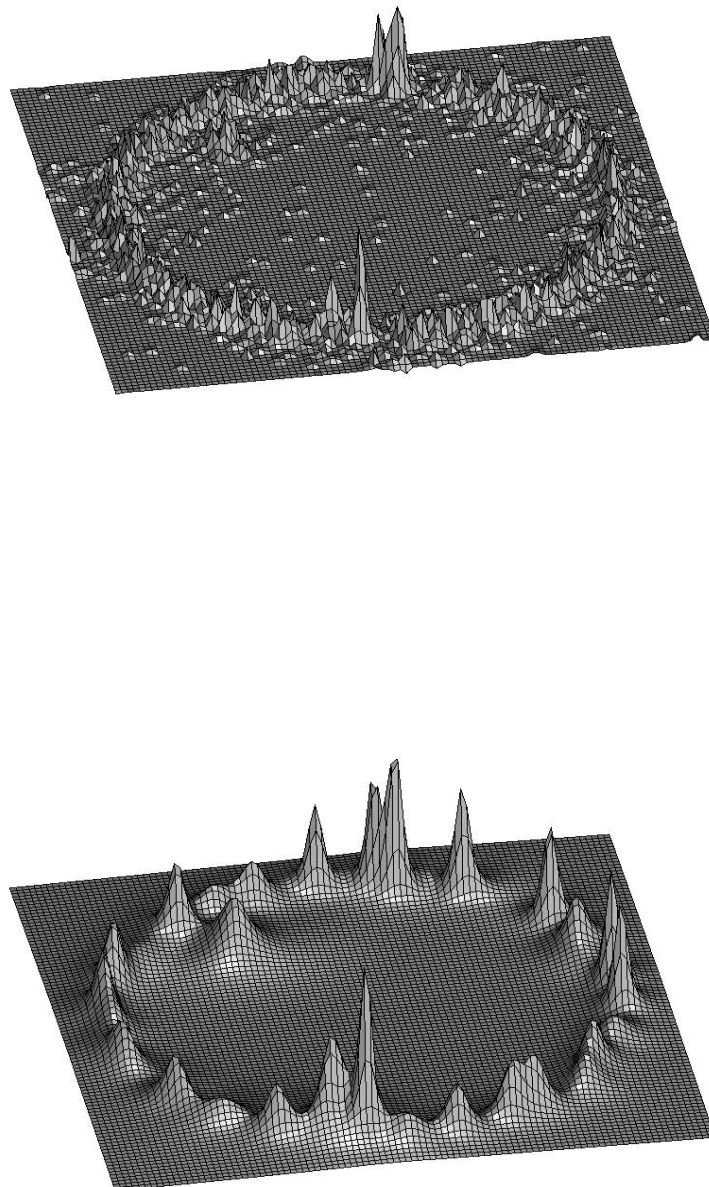


Figure 2: Top: Monte Carlo estimate of the condensed density when $\sigma = 0.8$; bottom: estimate of the condensed density by the closed form approximation with $\beta = 23.7$.

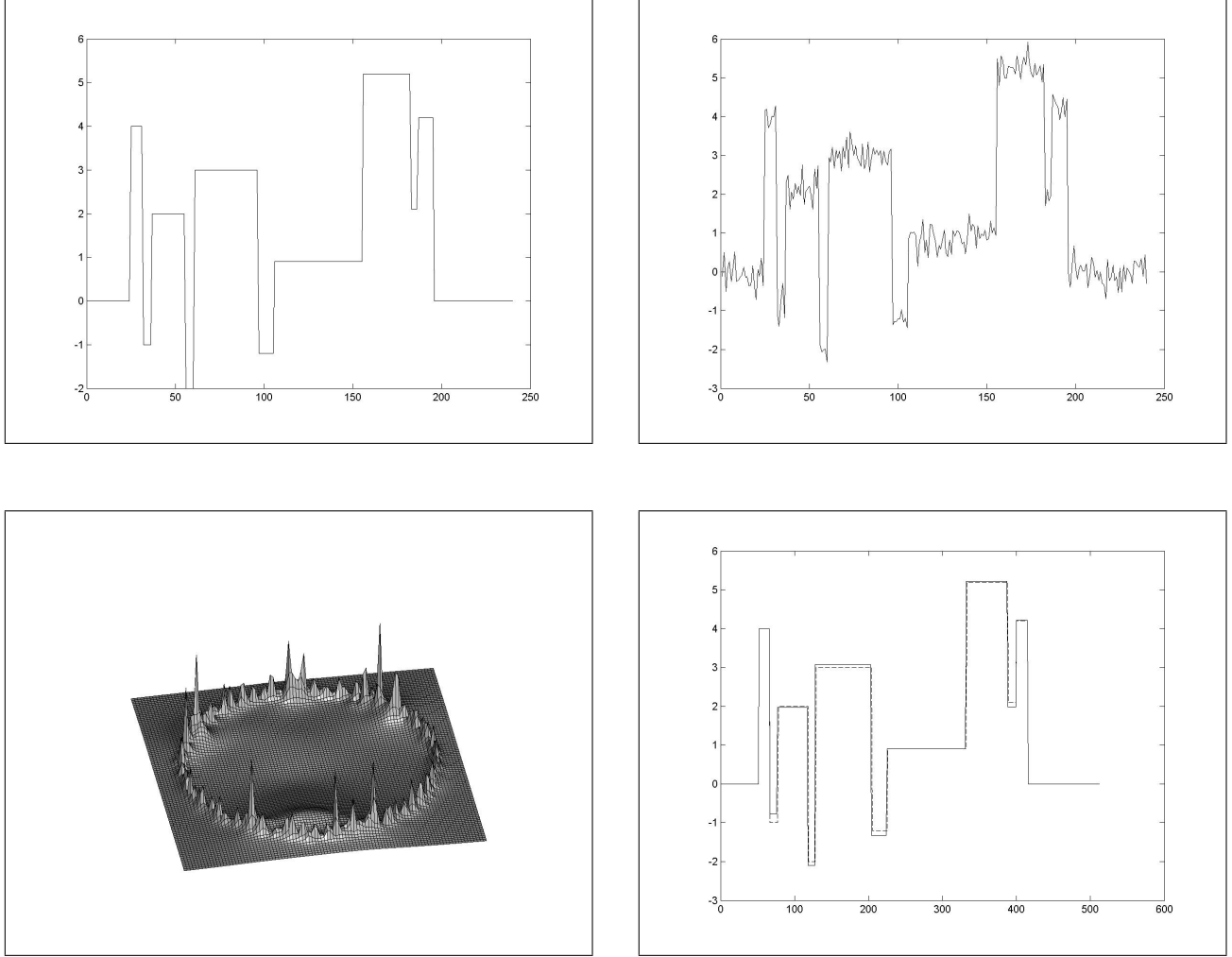


Figure 3: Top left: original function; top right: original function with Gaussian noise added (SNR=7). Bottom left: estimate of the condensed density by the closed form approximation; bottom right: reconstruction of the original function.

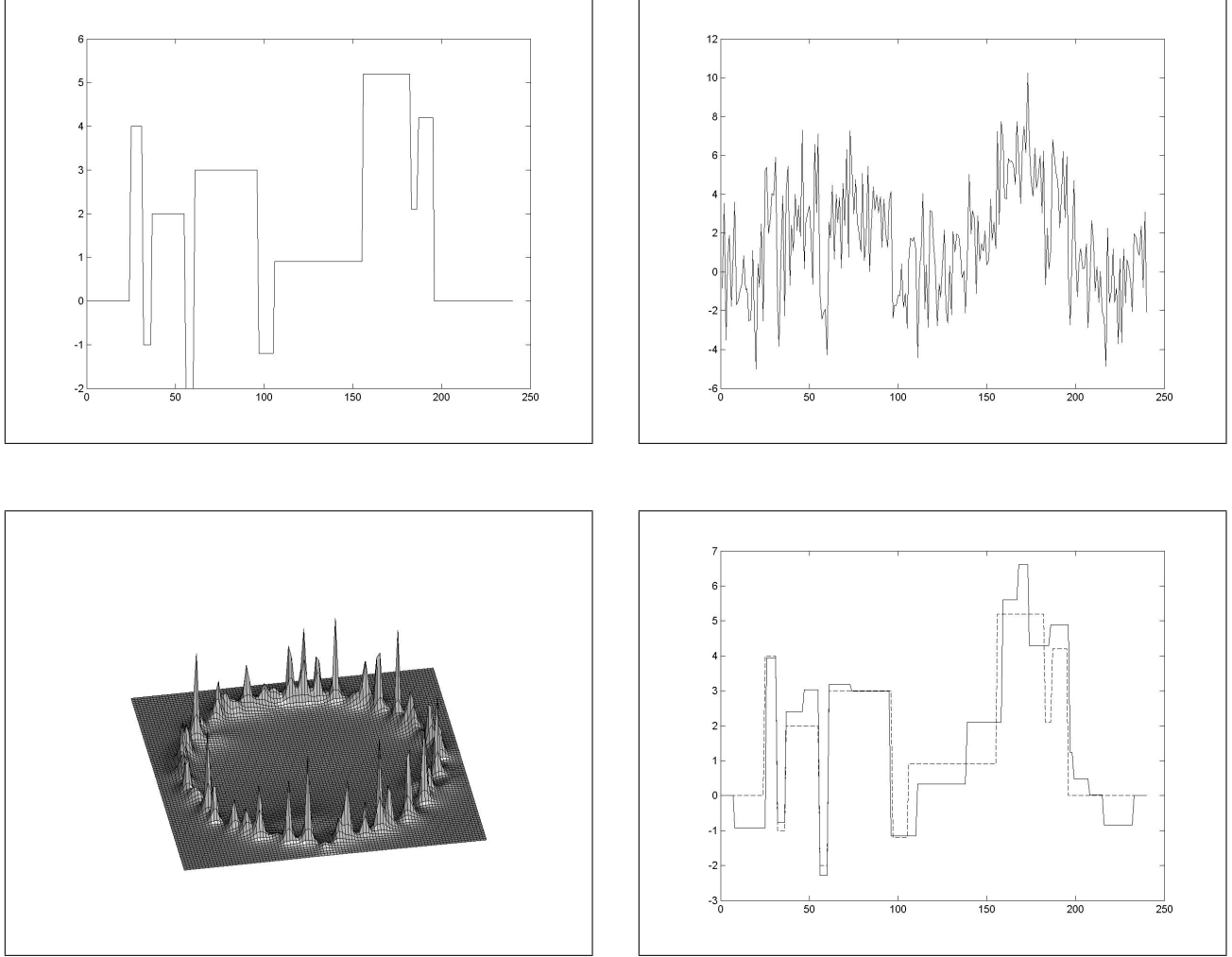


Figure 4: Top left: original function; top right: original function with Gaussian noise added (SNR=1). Bottom left: estimate of the condensed density by the closed form approximation; bottom right: reconstruction of the original function.